

WHAT IS CLAIMED IS:

1. In a system comprised of a plurality of client devices, each operating a browser application, and a host device operating a server application, a method for managing service by the server application, comprising:

receiving, by the server application from the browser applications, requests for services;

determining for each request a quality of service reflecting an ability of the server application to respond within a predetermined time based on a type associated with the request; and

returning, by the server application to at least one browser application, an appointment reflecting a time for the browser application to send another request to the server application for processing.

2. A load managing method comprising:

receiving by a server a request for service from a browser;

determining a quality of service for the request, the quality of service reflecting an ability of the server to respond to the request based on a type associated with the request; and

providing an appointment cookie to the browser reflecting a time at or after which the server will be available to process the request.

3. A load managing method comprising:

receiving by a server a request for service from a client;

determining a quality of service for the request, the quality of service reflecting an ability of the server to respond to the request based on a type associated with the request; and

providing an appointment to the client reflecting a time at or after which the server will be available to process the request.

4. A method of claim 3 wherein said type associated with a request is compared to prior requests of said type.

5. In a system having a client and a server, a load managing method comprising the steps, performed by the client, of:

transmitting a request for service to a server; and

receiving from the server an appointment cookie reflecting a time at which the server will be available to process the request.

6. The method of claim 5, further comprising:
enabling display of the appointment cookie.

7. The method of claim 5, further comprising:
re-transmitting the request to the server at or after the time reflected by the appointment cookie.

8. The method of claim 5, further comprising:

automatically without user intervention re-transmitting the request to the server at or after the time reflected by the appointment cookie.

9. In a system having a client and a server, a load managing method comprising the steps, performed by the client, of:

transmitting a request for service to a server; and

receiving from the server an appointment reflecting a time at which the server will be available to process the request.

10. In a system having a client and a server, a load managing method comprising the steps, performed by the client, of:

transmitting a request for service to a server;

receiving from the server an appointment reflecting a time at which the server will be available to process the request; and

resubmitting the request to the server application.

11. In a system having a client and a server, a load managing method comprising the steps, performed by the server, of:

receiving a request for service from the client;

determining a quality of service associated with providing the client with a response to the request; and

providing the client with an appointment reflecting a time at which the server will be available to process the request based on a result of the determination.

12. In a system having a client and a server, a load managing method comprising the steps, performed by the server, of:

receiving a request for service from the client;

determining a quality of service associated with providing the client with a response to the request; and

providing the client with an appointment cookie reflecting a time at which the server will be available to process the request based on a result of the determination.

13. A method for managing load on a process operating in a server computer, comprising:

receiving a request for a service to be performed by the process;

determining whether the process is able to perform the requested service subject to a predetermined level of quality; and

creating an appointment indicating that the process is not able to perform the requested service subject to a predetermined level of quality and reflecting a time at which it is determined that the process will be able to perform the requested service subject to a predetermined level of quality.

14. The method of claim 13, wherein the predetermined level of quality represents a time factor.

15. In a system comprising a first computer and a second computer, a method for managing service by a second computer, the method comprising:

- receiving, from the first computer, a request for service;
- determining for the request an estimated quality of service reflecting when the second computer will be able to respond to the request; and
- if the estimated quality of service is not satisfactory, returning, to the first computer, an appointment,

whereby the second computer will respond to the request for service at or after a time specified by the appointment.

16. In a system comprising a first computer and a second computer, a method for managing requests by a first computer, the method comprising:

- transmitting, to the second computer, a request for service;
- receiving an appointment, reflecting an estimated time when the second computer will be able to respond to the request; and
- resubmitting the request to the second computer on or after the time indicated in the appointment.

17. A system comprising:

- a server;
- a client;
- a network connecting the server to the client; and
- said server

- (i) receiving a request from said client,
- (ii) identifying a process capable of generating a response to the request,
- (iii) determining whether the identified process is able to generate the response subject to a predetermined level of quality,
- (iv) creating an appointment reflecting a time at which it is determined that the identified process will be able to generate the response subject to the predetermined level of quality, and
- (v) providing the appointment to the client.

18. The system of claim 17, wherein said client

- (i) receives the appointment from said server,
- (ii) determines based on the appointment when to send to said server a new request based on the request, and
- (iii) sends the new request to said server based on a result of the determination.

19. The system of claim 18, wherein said client

- (iv) includes in the new request information reflecting the appointment.

20. The system of claim 18, wherein said client

- (i) receives the appointment from said server,
- (ii) includes in the new request information reflecting the appointment.

21. A computer program product capable of configuring a data processor to manage service for a server application, the computer program product comprising program code to cause the data processor to perform the steps of:

receiving, by the server application from browser applications, requests for services;

determining for each request a quality of service reflecting an ability of the server application to respond within a predetermined time based on a type associated with the request; and

returning, by the server application to at least one browser application, an appointment reflecting a time for the browser application to send another request to the server application for processing.

22. A computer program product capable of configuring a data processor to manage loads, the computer program product comprising program code to cause the data processor to perform the steps of:

receiving by a server a request for service from a browser;

determining a quality of service for the request, the quality of service reflecting an ability of the server to respond to the request based on a type associated with the request; and

providing an appointment cookie to the browser reflecting a time at or after which the server will be available to process the request.

23. A computer program product capable of configuring a data processor to manage loads, the computer program product comprising program code to:

- receive by a server a request for service from a client;
- determine a quality of service for the request, the quality of service reflecting an ability of the server to respond to the request based on a type associated with the request; and
- provide an appointment to the client reflecting a time at or after which the server will be available to process the request.

24. A computer program product of claim 23 wherein said type associated with a request is compared to prior requests of said type.

25. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage loads, the computer program product comprising program code to:

- transmit a request for service to a server; and
- receive from the server an appointment cookie reflecting a time at which the server will be available to process the request.

26. The computer program product of claim 25, further comprising:
program code to enable display of the appointment cookie.

27. The computer program product of claim 25, further comprising:

program code to re-transmit the request to the server at or after the time reflected by the appointment cookie.

28. The computer program product of claim 25, further comprising:
program code to automatically without user intervention re-transmit the request to the server at or after the time reflected by the appointment cookie.

29. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage loads, the computer program product comprising program code to:

transmit a request for service to a server; and
receive from the server an appointment reflecting a time at which the server will be available to process the request.

30. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage loads, the computer program product comprising program code to:

transmit a request for service to a server;
receive from the server an appointment reflecting a time at which the server will be available to process the request; and
resubmit the request to the server application.

31. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage loads, the computer program product comprising program code:

receive a request for service from the client;

determine a quality of service associated with providing the client with a response to the request; and

provide the client with an appointment reflecting a time at which the server will be available to process the request based on a result of the determination.

32. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage loads, the computer program product comprising program code to cause the data processor to perform the steps of:

receive a request for service from the client;

determine a quality of service associated with providing the client with a response to the request; and

provide the client with an appointment cookie reflecting a time at which the server will be available to process the request based on a result of the determination.

33. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage loads, the computer program product comprising program code to:

receive a request for a service to be performed by the process;
determine whether the process is able to perform the requested service
subject to a predetermined level of quality; and
create an appointment indicating that the process is not able to perform the
requested service subject to a predetermined level of quality and reflecting a time at
which it is determined that the process will be able to perform the requested service
subject to a predetermined level of quality.

34. The computer program product of claim 33, wherein the predetermined
level of quality represents a time factor.

35. A computer program product, in a system having a client and a server,
capable of configuring a data processor to manage service, the computer program
product comprising program code to:

receive, from the first computer, a request for service;
determine for the request an estimated quality of service reflecting when the
second computer will be able to respond to the request; and
if the estimated quality of service is not satisfactory, return, to the first
computer, an appointment,
whereby the second computer will respond to the request for service at or
after a time specified by the appointment.

36. A computer program product, in a system having a client and a server, capable of configuring a data processor to manage service, the computer program product comprising program code to cause the data processor to perform the steps of:

transmit, to the second computer, a request for service;

receive an appointment, reflecting an estimated time when the second computer will be able to respond to the request; and

resubmit the request to the second computer on or after the time indicated in the appointment.

37. A system for managing services in a network having a plurality of client devices, each operating a browser application, and a host device operating a server application, the system comprising:

means for receiving, by the server application from the browser applications, requests for services;

means for determining for each request a quality of service reflecting an ability of the server application to respond within a predetermined time based on a type associated with the request; and

means for returning, by the server application to at least one browser application, an appointment reflecting a time for the browser application to send another request to the server application for processing.

38. A system for load management comprising:

means for receiving by a server a request for service from a browser;

means for determining a quality of service for the request, the quality of service reflecting an ability of the server to respond to the request based on a type associated with the request; and

means for providing an appointment cookie to the browser reflecting a time at or after which the server will be available to process the request.

39. A system for load management comprising:

means for receiving by a server a request for service from a client;

means for determining a quality of service for the request, the quality of service reflecting an ability of the server to respond to the request based on a type associated with the request; and

means for providing an appointment to the client reflecting a time at or after which the server will be available to process the request.

40. A system of claim 39 wherein said type associated with a request is compared to prior requests of said type.

41. A system for load management having a client and a server, the system comprising the steps, performed by the client, of:

means for transmitting a request for service to a server; and

means for receiving from the server an appointment cookie reflecting a time at which the server will be available to process the request.

Page 49 of 50

42. The system of claim 41, further comprising:
means for enabling display of the appointment cookie.

43. The system of claim 41 further comprising:
means for re-transmitting the request to the server at or after the time
reflected by the appointment cookie.

44. The system of claim 41, further comprising:
means for automatically without user intervention re-transmitting the request
to the server at or after the time reflected by the appointment cookie.

45. A system for load management having a client and a server, the
system comprising the steps, performed by the client, of:
means for transmitting a request for service to a server; and
means for receiving from the server an appointment reflecting a time at which
the server will be available to process the request.

46. A system for load management having a client and a server, the
system comprising the steps, performed by the client, of:
means for transmitting a request for service to a server;
means for receiving from the server an appointment reflecting a time at which
the server will be available to process the request; and

means for resubmitting the request to the server application.

47. A system for load management having a client and a server, the system comprising the steps, performed by the server, of:

means for receiving a request for service from the client;

means for determining a quality of service associated with providing the client with a response to the request; and

means for providing the client with an appointment reflecting a time at which the server will be available to process the request based on a result of the determination.

48. A system for load management having a client and a server, the system comprising the steps, performed by the server, of:

means for receiving a request for service from the client;

means for determining a quality of service associated with providing the client with a response to the request; and

means for providing the client with an appointment cookie reflecting a time at which the server will be available to process the request based on a result of the determination.

49. A system for load management a process operating in a server computer, comprising:

means for receiving a request for a service to be performed by the process;

202-408-4000

means for determining whether the process is able to perform the requested service subject to a predetermined level of quality; and

means for creating an appointment indicating that the process is not able to perform the requested service subject to a predetermined level of quality and reflecting a time at which it is determined that the process will be able to perform the requested service subject to a predetermined level of quality.

50. The system of claim 49, wherein the predetermined level of quality represents a time factor.

51. In a system for managing services comprising a first computer and a second computer, a system for managing service by a second computer, the system comprising:

means for receiving, from the first computer, a request for service;

means for determining for the request an estimated quality of service reflecting when the second computer will be able to respond to the request; and

if the estimated quality of service is not satisfactory, means for returning, to the first computer, an appointment,

whereby the second computer will respond to the request for service at or after a time specified by the appointment.

means for resubmitting the request to the second computer on or after the time indicated in the appointment.

36